Abstract

A method for synchronising data between a receiving computer and a sending computer by way of a computer network or the like, wherein the sending computer has a source file and the receiving computer has a reference file. The source file is arranged at the sending computer into a sequence of data blocks, each block comprising a predetermined number of data units (e.g. bytes), and a source key value is computed for each block in the source file, and the source key values are transmitted to the receiving computer. At the receiving computer, the source key values are compared with reference key values computed for each predetermined number of contiguous data units in the reference file to determine matches between source key values and reference key values. Then, based on indications of which source keys do not have matching reference keys, specific data blocks are transmitted from the source file corresponding to the unmatched source keys. Then, the receiving computer constructs a target file from the contiguous data units in the reference file determined to have reference key values matching respective source key values and the data blocks from the source file received from the sending computer. In the context of the Internet, the provision of the source key values and data blocks to the receiving computer can conveniently be carried out in appropriate circumstances by a proxy server or the like, using standard caching techniques.